ARO 17339.3-MA

48

(2)

BEST AVAILABLE COPY

AD A126539

Remote Sensing of Temperature Profiles in the Atmosphere

by

Finbarr O'Sullivan
Department of Statistics
University of Wisconsin
1210 W. Dayton Street
Madison, WI 53706

April 1982

proceedings

Presented at the NASA workshop on "Density estimation and function smoothing" held at Texas A&M March 11-13, 1982.

This research was supported by the USARO under Grant No. DAAG29-80-K0042.

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER 17339.3-MA	2. GOVT ACCESSION NO. N/A	3. RECIPIENT'S CATALOG NUMBER N/A
4. TITLE (and Subtitle) Remote Sensing of Temperature Profiles in the Atmosphere		5. TYPE OF REPORT & PERIOD COVERED Reprint
		6. PERFORMING ORG. REPORT NUMBER N/A
7. AUTHOR(s) Finbarr O'Sullivan		8. CONTRACT OR GRANT NUMBER(s) DAAG29 80 K 0042
9. PERFORMING ORGANIZATION NAME AND ADDRESS University of Wisconsin Madison, WI 53706		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS N/A
11. CONTROLLING OFFICE NAME AND ADDRESS U. S. Army Research Office P. O. Box 12211 Research Triangle Park, NC 27709		12. REPORT DATE 1982
		13. NUMBER OF PAGES 12
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Submitted for announcement only.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

83 04 07 080

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

1. Introduction

Remote sensing of the atmosphere is a rapidly developing science. Today's meteorological satellites such as those in the TIROS-N series have high resolution instruments on board which measure the intensity of upwelling radiation in selected channel frequencies. A description of the data retrieved by the radiometers on the TIROS-N type satellites can be found in [7]. From these data it is possible to obtain information on the atmosphere's temperature, moisture and wind structure. One of the goals of the current Satellite Meteorology program is to improve the quality of atmospheric information obtained from satellite soundings to a point where it can be used for weather forecasting purposes. A major challenge in this direction is to develop refined numerical and statistical methods for inverting the equations of radiative transfer given a finite number of noisy measurements.

For a non-scattering atmosphere in local thermodynamic equilibrium the radiative transfer equations (RTE's) describe how the satellite upwelling radiance measurements relate to the underlying temperature distribution T:-

$$\mathcal{I}_\nu(T) = B_\nu[T(p_0)]\tau_\nu(p_0) - \int_0^{p_0} B_\nu[T(p)\frac{d}{dp}\tau_\nu(p)dp \qquad (1.1)$$

where p_0 is the surface pressure, $\tau_\nu(p)$ is the transmittance of the atmosphere above pressure p at frequency ν, and B_ν is Plank's function given by:-

$$B_\nu[T(p)] = c_1\nu^3/\{\exp(c_2\nu/T(p)) - 1\}$$
$$c_1 = 1.19061\text{X}10^{-5}\text{erg-cm}^2\text{-sec}^{-1} \quad (1.2)$$
$$c_2 = 1.43868 \text{ cm-deg(K)}$$

The R.T.E's are of course an idealization. They describe the intensities the satellite radiometer would record in the absence of such things as atmospheric attenuation due to clouds or instrument noise. However, by using high resolution radiometers like the HIRS or AVHRR, sets of intensity measurements from many FOV's (fields of vision) can be combined to obtain data of the form

$$z_i = \mathcal{J}_{\nu_i}(T) + e_i \qquad i = 1,\ldots,n \quad (1.3)$$

where e_i's are errors. These data relate to an area of about 119 by 140 km on the earth's surface. See [6] for more details.

We are interested in refining the methods used to obtain temperature distribution estimates from the above data. The procedure currently used to process TIROS-N temperature sounding data is a linear regression technique see [6]. Here we begin to discuss how the method of regularization (M.O.R.) might be used to improve the quality of temperature profiles obtainable by this procedure.

Let T be the true temperature profile in the atmosphere. Then T can be written as

$$T = T_0 + \delta \tag{1.4}$$

where T_0 is the current best guess of T and δ is the update or correction to T_0 to be estimated from the data $\{z_i\}$ in hand. Using M.O.R. to estimate δ involves consideration of a functional I_λ given by

$$I_\lambda(\delta) = \sum_{i=1}^{n}[z_i-\mathcal{J}_\nu(T_0+\delta)]^2 + \lambda \int_0^{p_0}[\delta^{(m)}(p)]^2 dp \tag{1.5}$$

and picking the estimated update δ_λ to minimize this functional[1] over some class of physically plausible candidates, for instance the set of functions δ in $W_2^m[0,p_0]$ for which $T_0+\delta$ is positive or perhaps, if the location of the temperature inversion were reliably known, one would look for minimizers of I_λ subject to an additional constraint involving temperature inversion.

The statistical reasoning for considering regularized estimates of this type is well documented in the literature, see for example [3] and [1]. Intuitively δ_λ has been designed to match the observed data and possess certain smoothness qualities. The parameter λ controls a tradeoff between the smoothness of a solution (measured by $\int_0^{p_0}[\delta_\lambda^{(m)}(p)]^2 dp$) and how well it

---

[1] This corresponds to the case when the measurement errors are iid $N(0,\sigma^2)$. A more "robust" method would be to consider functionals of the form

$$I_\lambda(\delta) = \sum_{i=1}^{n}\rho[z_i-\mathcal{J}_{\nu_i}(T_0+\delta)] + \lambda \int_0^{p_0}[\delta^{(m)}(p)]^2 dp$$

where ρ reflected the possible non-Gaussian nature of the noise.

matches the data (the $\sum_{i=1}^{n}[z_i-\mathcal{J}_{\nu_i}(T_0+\delta_\lambda)]^2$ term).

Inverting the R.T.E.'s with noisy data can be viewed as a special case of a more general situation in which the scientist wishes to estimate a function x given data

$$z_i = N_i(x) + e_i \quad i = 1,\ldots,n \tag{1.6}$$

where x is in some Hilbert space H, the N_i's are non-linear functionals and e_i's are noise. Here, assuming the e_i's are iid $N(0,\sigma^2)$, an appropriate regularization function I_λ is

$$I_\lambda(x) = \sum_{i=1}^{n}[z_i-N_i(x)]^2 + \lambda J(x) \tag{1.7}$$

where J is a roughness penalty functional on H. To estimate x one proceeds to minimize I_λ over some subset of physical interest in H. This report summarizes recent results we have obtained on the existence and numerical approximability of minimizers of such I_λ's in certain subsets of H. We indicate how these results apply to the radiative transfer equations case.

There are three sections: section 2 talks about the existence theory; a Gauss-Newton algorithm for minimizing the regularization functionals is outlined in section 3, while the final section briefly describes how to estimate the smoothing parameter using a first order approximation to the generalized cross validation function given in [8]. We assume the reader is familiar with the basic mathematical tools for discussing minimization problems in Hilbert spaces. Part 1 of Ekeland and Temam's book [2] is an inspiring introduction to this subject.

2. Existence Theory

Preliminaries

Before describing our main results, let's pause a moment to get our notation straight. H is a real Hilbert space with inner product $\langle\cdot,\cdot\rangle$ and norm $||\cdot||$ (so $\langle x,x\rangle=||x||^2$). P is a projection operator in H with finite dimentional null space; the complementary projection I-P is denoted by P_0. H^* is the dual space of H, i.e. the space of all continuous linear maps from H into R. $L(H,H^*)$ is the space of linear operators from H into H^*. We will discuss functionals, I say, acting on H (so $I: H \to R$). The first and second Frechet derivatives of I at some point $x \in H$ will be denoted by $I'(x)$ and $I''(x)$ respectively. Think of $I'(x)$ as an element of H^* and $I''(x)$ as an element of $L(H,H^*)$. Our concern here is with regularization functionals I_λ on H given by

$$I_\lambda(x) = \sum_{i=1}^{n} [z_i - N_i(x)]^2 + \lambda ||Px||^2 \tag{2.1}$$

where N_i's are functionals on H, z_i's are in $\mathbb{R}$, $x \in H$ and $\lambda > 0$. Whenever we write I_λ the form (2.1) will be what is meant. So we are considering regularization procedures in which the roughness penalty $J(x)$ is a semi-norm on H given by $J(x) = ||Px||^2$.

Main Results

We now specify conditions on the non-linear functionals N_i which guarantee the existence of minimizers of I_λ in closed convex subsets K of H. In the R.T.E. case a reasonable choice for K is the set of all functions in $W_2^m[0,p_0]$ for which $T_0+\delta$ is positive. It is very easy to check that this K is a closed convex subset of $W_2^m[0,p_0]$ for any m. Our existence results are summarized in the following three theorems.

Theorem 1 (proof in [2] pp. 34-35).

Let K be a closed convex subset of a Hilbert space H. Suppose $I_\lambda: K \to R$ is coercive on K (i.e. $I_\lambda(x) \to \infty$ as $||x|| \to \infty$ in K) and moreover that I_λ is weakly lower semi-continuous (w.l.s.c.) on K then I_λ attains its infimum on K.

Theorem 2 (proof in [4])

Let $\phi: R \to R$ be a monotonic increasing function in the modulus of its argument. Suppose

(i) $\sum_{i=1}^{n} \phi(N_i(x))$ is convex on K

(ii) $\sum_{i=1}^{n} \phi[N_i(x)] = \phi \Leftrightarrow P_0 x = P_0 \theta$ for some θ in K

then I_λ is coercive on K.

Remark: The above theorem can be generalized somewhat but we refrain from doing so because the form given has more intuitive appeal.

Theorem 3

If N_i is weakly continuous (w.c.) on K for each i then I_λ is w.l.s.c. on K.

Proof: If the N_i are w.c., then it surely follows that $\sum_{i=1}^{n} [z_i - N_i(x)]^2$ is w.c. But $||Px||^2$ is well known to be w.l.s.c. Therefore I_λ is w.l.s.c. QED

Application to the R.T.E.'s (see [4] for details)

The $\mathcal{J}_{\nu_i}$ arising here can be shown to satisfy the hypotheses of Theorem 2 with ϕ taken to be

$$\phi(x) = |x|, \ x\varepsilon R$$

Also, each $\mathcal{J}_{\nu_i}$ is w.c. We therefore have that for each $\lambda>0, \exists\ \delta_\lambda \varepsilon K = \{\delta\varepsilon W_2{}^m[0,p_0]\ T_0+\delta>0\}$, s.t.

$$I_\lambda(\delta_\lambda) = \min_{\delta\varepsilon K} \{ \sum_{i=1}^{n} [z_i - \mathcal{J}_{\nu_i}(T_0+\delta)]^2 + \lambda\int_0^{P_0} [\delta^{(m)}(P)]^2 dp\}$$

There exist regularized solutions to the R.T.E.'s.

3. A numerical procedure for minimizing I_λ in K

Let x^k be the k^{th} approximation to the minimizer in K of I_λ. Define the functional $I_\lambda{}^k$ on K as follows

$$I_\lambda{}^k(x) = \sum_{i=1}^{n} [z_i - N_i(x^k) - N_i'(x^k)[x-x^k]]^2 + \lambda||Px||^2 \qquad (3.1)$$

each N_i is simply linearized about x^k. Define x^{k+1} to be the minimizer in K of $I_\lambda{}^k$.

Under suitable regularity conditions the iterates x^k are well defined and can be shown to satisfy

$$\begin{aligned} x^{k+1} &= x^k - \{\sum_{i=1}^{n} N_i'(x^k)N_i'(x^k) + \lambda\langle P\cdot,\cdot\rangle\}^{-1} I_\lambda'(x^k) \\ &\equiv x^k - A^{-1}(x^k)I'(x^k) \end{aligned} \qquad (3.2)$$

That this equation makes good sense is evident once one realizes that $A(x^k)$ belongs to $L(H,H^*)$ and $I_\lambda'(x^k)$ is in H^*.

Those in the know will have recognized that the above procedure is nothing more than an infinite dimensional version of the Gauss-Newton algorithm. The finite dimentional case is discussed in [5]. The major advantage of using a Gauss-Newton procedure to minimize our regularization functionals is the ease with which successive interates can be obtained. At each stage we have a regularization problem involving linear functionals, the $N_i'(x^k)$'s, consequently we can take advantage of available software tools.

With the appropriate assumptions it is possible to show that the procedure is a decent method and the sequence x^k converges at least R-linearly to a critical point of I_λ in K.

<u>Theorem 4</u> (proof in [4]).

Suppose that the $N_i(.)$'s are twice continuously differentiable and $N_i'(.)$'s are w.c. on int K. Let $x^o \varepsilon$ int K be such that

$$L^o = \{x | I_\lambda(x) < I_\lambda(x^o)\}$$

is weakly compact and I_λ has only finitely many critical points in L^o. Moreover, suppose that μ_0, μ_1, γ_1 all positive with $\mu_0 - 1/2\gamma_1 > 0$ satisfying

$$\mu_0 ||h||^2 < \langle h, A(x)h \rangle < \mu_1 ||h||^2, \; I_\lambda''(x)hh < \gamma_1 ||h||^2 \quad \forall x \varepsilon L^o, \; h \varepsilon H$$

then the sequence of iterates $\{x^k\} \subseteq L^o$, $\lim_k x^k = x^*$ where $I_\lambda'(x^*) \equiv 0$ and if $I_\lambda''(x^*)$ is non-singular, then the convergence is at least R-linear. The proof follows an argument similar to that used in 14.4.6 of [5].

4. The choice of λ

The generalized cross validation method for choosing λ works as follows. Let $x_\lambda^{[k]}$ be the minimizer[2] in K of

$$\sum_{\substack{i=1\\ i\neq k}}^{n} [z_i - N_i(x)]^2 + \lambda ||Px||^2 \tag{4.1}$$

Then λ is chosen to minimize

$$V(\lambda) = \frac{\frac{1}{n}\sum_{k=1}^{n} [z_k - N_k(x_\lambda^{[k]})]^2}{[1-\frac{1}{n}\sum_{k=1}^{n} a_{kk}^*(\lambda)]^2}$$

where $N_k(x_\lambda^{[k]})$ is the prediction of z_k given the data $z_1, z_2, \ldots, z_{k-1}$ $z_{k+1} \ldots z_n$ and $a_{kk}^*(\lambda)$ is the "differential influence" of the z_k'th data point on the estimate x_λ (x_λ is the minimizer in K of I_λ).

$$a_{kk}^*(\lambda) = \frac{N_k(x_\lambda^{[k]}) - N_k(x_\lambda)}{N_k(x_\lambda^{[k]}) - z_k} \tag{4.3}$$

From a computational viewpoint $V(\lambda)$ is prohibitively expensive so one needs to find some convenient approximation. Following Wahba [8], $V(\lambda)$ can be approximated by

$$V_{approx}(\lambda) = \frac{\frac{1}{n}\sum_{k=1}^{n} [z_k - N_k(x_\lambda)]^2}{[1-\mu_1(\lambda)]^2} \tag{4.4}$$

[2] Assumed to be uniquely defined.

where μ_1 given by

$$\mu_1(\lambda) = \frac{1}{n}\sum_{k=1}^{n} \frac{\partial N_k(x_\lambda)}{\partial z_k}$$

is an easily computed functional of x_λ. We hope to study this procedure more closely in the near future.

Acknowledgements

This work was done with the help of my thesis advisor, Professor Grace Wahba.

REFERENCES

[1] D.D. Cox, "Asymptotics for M-Type Smoothing Splines," Statistics Department, University of Wisconsin, Madison, Technical Report No. 654, November 1981.

[2] J. Ekeland and R. Teman, Analyse Convexe et Problems Variationelles, Herman, Paris (1973).

[3] G.S. Kimeldorf and G. Wahba, "A correspondence between Bayesian Estimation in Stochastic Processes and Smoothing by Splines," Annals of Mathematical Statistics Al, pp. 495-562 (1970).

[4] F. O'Sullivan, Thesis to appear. (1982)

[5] J.M. Ortega and W.C. Rheinbold, Iterative Solutions of Non-linear Equations to Several Variables, Academic Press (1970).

[6] W.L. Smith and H. M. Woolf, "The Use of Eigenvectors of Statistical Covariance Matrices for Interpreting Sastellite Sounding Radiometer Observations," Journal of the Atmospheric Sciences 33, 7, pp. 1127-1140, July 1976.

[7] W.L. Smith, H.M. Woolf, C.M. Hayden, D.O. Wark, and L.M. McMillin, "The TIROS-N Operational Vertical Sounder," Bulletin of the American Meteorological Society 50, 10, pp. 1177-1187, October 1979.

[8] G. Wahba, "Constrained Regularization for Ill-Posed Linear Operator Equations, with Applications in Meteorology and Medicine," Statistics Department, University of Wisconsin, Madison, Technical Report No. 646, August 1981.